Available online at www.sciencedirect.com

SCIENCE DIRECT®

LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER

Linear Algebra and its Applications 389 (2004) 189–213

www.elsevier.com/locate/laa

New vector sequence transformations

C. Brezinski [a,*,1], M. Redivo-Zaglia [b]

[a] *Laboratoire Paul Painlevé, UMR CNRS 8524, UFR de Mathématiques Pures et Appliquées, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq cedex, France*

[b] *Dipartimento di Matematica Pura ed Applicata, Università degli Studi di Padova, Via G.B. Belzoni 7, 35131 Padova, Italy*

Received 2 April 2003; accepted 19 March 2004

Submitted by R.A. Brualdi

Abstract

Many numerical methods produce sequences of vectors converging to the solution of a problem. When the convergence is slow, the sequence can be transformed into a new vector sequence which, under some assumptions, converges faster to the same limit. The construction of a sequence transformation is based on its kernel, that is the set of sequences which are transformed into a constant sequence. In this paper, new vector sequence transformations are built from kernels which extent those of the most general transformations known so far.

AMS classification: 65B05; 65F99

Keywords: Sequence transformation; Extrapolation; Convergence acceleration; Vector sequences; Schur complement

1. Introduction

Let (x_n) be a sequence of vectors in $\mathbb{R}^p$ converging to $x \in \mathbb{R}^p$ when n goes to infinity. If this sequence converges slowly, one can try to accelerate its convergence by transforming it into a new sequence (y_n) converging faster to the same limit. Such

* Corresponding author. Tel.: +33-3-2043-4296; fax: +33-3-2043-6869.

E-mail addresses: claude.brezinski@univ-lille1.fr (C. Brezinski), michela.redivozaglia@unipd.it (M. Redivo-Zaglia).

[1] The work of C. Brezinski was partially supported by the INTAS Research Network NeCCA (INTAS 03-51-6637).

0024-3795/$ - see front matter © 2004 Published by Elsevier Inc.
doi:10.1016/j.laa.2004.03.031

a procedure is called a (*vector*) *sequence transformation*. Each sequence transformation $T : (x_n) \mapsto (y_n)$ is built in such a way that $\forall n,\ y_n = x$ when applied to a test sequence (x_n) satisfying a certain relationship depending on an unknown vector x (usually the limit of (x_n)) and on unknown parameters. The set of such test sequences is called the *kernel* of T. Then, if the transformation T is applied to a sequence (x_n) not too different from a sequence of its kernel, it is hoped that the sequence (y_n) it produces will converge to x faster than (x_n).

The most general kernel considered so far consists of vector sequences of the form

$$x_n = x + a_1 z_n^1 + \cdots + a_k z_n^k, \quad n = 0, 1, \ldots, \tag{1}$$

where the a_is are unknown scalars, x is an unknown vector and the (z_n^i)s are known sequences of vectors in $\mathbb{R}^p$ for $i = 1, \ldots, k$. This case will be called the *unknown scalar case* since the unknowns a_i are scalars.

Let us explain how to construct a transformation T with this kernel. If the scalars $a_1, \ldots, a_k$ in (1) are known, and if we define the transformation $T : (x_n) \mapsto (y_n)$ by

$$y_n = x_n - a_1 z_n^1 - \cdots - a_k z_n^k, \quad n = 0, 1, \ldots, \tag{2}$$

then, by construction, we will have $\forall n,\ y_n = x$.

Several strategies for the computation of $a_1, \ldots, a_k$ were studied in [5,8,9,12] and [41, pp. 175–179]. They lead to various vector sequence transformations which can be implemented by different recursive algorithms. They all consist in using (1) for writing down k linear algebraic equations involving the unknowns $a_1, \ldots, a_k$, solving this system, and then computing y_n by (2). It is also possible to write down a system of $k + 1$ equations in the unknowns $y_n, a_1, \ldots, a_k$ and to solve it for y_n.

Applying the transformation T to a sequence (x_n) which does not belong to its kernel is equivalent to interpolating it by a sequence of the kernel, then computing the unknowns a_i as above (they now depend on n and, of course, on k), and finally defining y_n by (2). This is the reason why a sequence transformation is often called an *extrapolation method*.

A more general case, where the kernel is given by

$$x_n = x + Az_n, \quad n = 0, 1, \ldots,$$

where A is an unknown $p \times q$ matrix and (z_n) a known sequence of vectors in $\mathbb{R}^q$, was considered in [15].

In this paper, we consider two even more general kernels and we construct the corresponding vector sequence transformations. Direct formulations and recursive algorithms will be obtained. They have a form quite similar to vector sequence transformations such as the E-transformation (and the E-algorithm for its implementation) [5], and the RPA and its variants [6]. However, although some of these algorithms have the same kernel, they correspond to different sequence transformations. Transformations in the style of those given in [8] could also be constructed.

General references on sequence transformations and extrapolation algorithms are [10,41].

2. Mathematical ingredients

In this paper, we will make use of the notions of pseudo-inverse of a matrix, pseudo-Schur complements, and bordered matrices. Let us remind the definitions and the properties which are the most important for our purpose.

2.1. Pseudo-inverses

The pseudo-inverse (Moore–Penrose inverse) of a matrix $A \in \mathbb{R}^{p\times q}$ is the unique matrix $A^\dagger \in \mathbb{R}^{q\times p}$ satisfying

$$\begin{aligned} A^\dagger A A^\dagger &= A^\dagger, \\ A A^\dagger A &= A, \\ (A^\dagger A)^{\mathrm{T}} &= A^\dagger A, \\ (A A^\dagger)^{\mathrm{T}} &= A A^\dagger. \end{aligned}$$

We have $(A^\dagger)^\dagger = A$, $(A^\dagger)^{\mathrm{T}} = (A^{\mathrm{T}})^\dagger$ and $(A^{\mathrm{T}}A)^\dagger = A^\dagger(A^\dagger)^{\mathrm{T}}$.

If only some of these conditions are satisfied, the corresponding matrix $A^\dagger$ is often called a generalized inverse [3,36]. If A is square and nonsingular, the pseudo-inverse and the generalized inverse coincide with the inverse.

There are two special cases, which will be useful later, where the expression of the pseudo-inverse simplifies. They are both useful in the solution of the system of linear equations $Ax = b$ in the least squares sense.

(1) If $p \geqslant q$ and $\mathrm{rank}(A) = q$, then
$$A^\dagger = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}},$$
and we have $A^\dagger A = I \in \mathbb{R}^{q\times q}$.
In that case, a least squares solution of the overdetermined system $Ax = b$ satisfies $A^{\mathrm{T}}Ax = A^{\mathrm{T}}b$ and thus $x = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b = A^\dagger b$.

(2) If $p \leqslant q$ and $\mathrm{rank}(A) = p$, then
$$A^\dagger = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}$$
and it holds $AA^\dagger = I \in \mathbb{R}^{p\times p}$.
In that case, the system $Ax = b$ is underdetermined, it is consistent, and its minimum Euclidean norm solution is obtained by setting $x = A^{\mathrm{T}}y$. Thus the system writes $AA^{\mathrm{T}}y = b$, which gives $y = (AA^{\mathrm{T}})^{-1}b$ and it follows $x = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b = A^\dagger b$.

So, in both cases, we consider the rectangular system of linear equations $Ax = b$ where $A \in \mathbb{R}^{p\times q}$ has rank $k \leqslant \min(p, q)$, $x \in \mathbb{R}^q$ and $b \in \mathbb{R}^p$. The least squares solution of the problem of finding

$$\min_{x\in V} \|x\|_2, \quad V = \{x \in \mathbb{R}^q \mid \|Ax - b\|_2 = \min\}$$

is given by $x = A^{\dagger}b$. Similarly, a system with several right hand sides $AX = B$, where $X \in \mathbb{R}^{q\times m}$ and $B \in \mathbb{R}^{p\times m}$, could be treated.

It follows from these two particular cases that, if $A \in \mathbb{R}^{p\times q}$ and $B \in \mathbb{R}^{q\times m}$, with $p \geqslant q$ and $q \leqslant m$, and if both matrices have rank q then [4, p. 17]

$$(AB)^{\dagger} = B^{\dagger}A^{\dagger} = B^{\mathrm{T}}(BB^{\mathrm{T}})^{-1}(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}.$$

Necessary and sufficient conditions for the property $(AB)^{\dagger} = B^{\dagger}A^{\dagger}$ to hold are given in [27].

For more results on pseudo-inverses, see [4] or [20].

Although, in this paper, we employ pseudo-inverses, their four properties are not always used for each of our transformations. In fact, for some of them, we only need the left inverse X^{L} of a matrix X which satisfies $X^{\mathrm{L}}X = I$. In particular, X^{L} can be defined by $X^{\mathrm{L}} = (Y^{\mathrm{T}}X)^{-1}Y^{\mathrm{T}}$, where Y is a matrix of the same dimensions as X and such that $Y^{\mathrm{T}}X$ is nonsingular. Right inverses could be defined similarly.

2.2. *Pseudo-Schur complements*

A subject related to pseudo-inverses is the Schur complement and its generalizations [25,40,42]. This important notion was introduced by Schur [40] in 1917 (see, for example, [25, pp. 19–23] or [42, Section 6.4] for basic definitions and results). Schur complements have many applications in extrapolation methods and related topics [7,8,13] and in other domains of mathematics, in particular in linear algebra [23,24,26,28,34].

We consider the matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{3}$$

where $A \in \mathbb{R}^{p\times q}$, $B \in \mathbb{R}^{p\times s}$, $C \in \mathbb{R}^{r\times q}$ and $D \in \mathbb{R}^{r\times s}$. Thus $M \in \mathbb{R}^{(p+r)\times(q+s)}$.

If $r = s$ and if D is nonsingular, the Schur complement of D in M is denoted by (M/D) and it is defined by

$$(M/D) = A - BD^{-1}C.$$

Moreover, if $p = q$, the Schur determinantal formula holds

$$\det(M/D) = \frac{\det M}{\det D}. \tag{4}$$

In [6], this formula was extended to the case where $q = 1$ and $r = s$ by defining $\det M$ as the vector of $\mathbb{R}^p$ which is the combination of the vector A and of the columns of B which form its first row and is obtained by the classical rule for expanding a determinant (if needed, see [8] for more explanations). With this definition, it follows that $\det(M/D) = (M/D)$. In the sequel, this formula will be called the *vector Schur determinantal formula*.

Let us now define the pseudo-Schur complement of D in M, denoted by $(M/D)_{\mathscr{P}}$, by

$$(M/D)_{\mathscr{P}} = A - BD^{\dagger}C, \tag{5}$$

where $D^{\dagger}$ is the pseudo-inverse of D. If D is square and nonsingular, the usual Schur complement is recovered. Pseudo-Schur complements were introduced in [22,32,33], but they were also implicitly considered in [2,36]; see [1,19,21] for reviews. Pseudo-Schur complements of A, B and C in M can be defined similarly. A determinantal formula similar to Schur's does not seem to hold for the pseudo-Schur complement in the general case.

If $q = 1$, A is a vector in $\mathbb{R}^p$ and we recover the vector Schur determinantal formula.

A matrix pseudo-Schur complement can be defined from the vector pseudo-Schur complement considered above.

More results on pseudo-Schur complements are given in [35].

2.3. Bordered matrices

We will now discuss the pseudo-inverse of the bordered matrix M. We set

$$M^{\ddagger} = \begin{pmatrix} A^{\dagger} + A^{\dagger}BS^{\dagger}CA^{\dagger} & -A^{\dagger}BS^{\dagger} \\ -S^{\dagger}CA^{\dagger} & S^{\dagger} \end{pmatrix}, \tag{6}$$

where $S = D - CA^{\dagger}B \in \mathbb{R}^{r\times s}$ is the pseudo-Schur complement of A in M.

Such an expression was already considered in [2]. It must be noticed that the matrix $M^{\ddagger}$ defined by (6) is not, in general, the pseudo-inverse of M. Necessary and sufficient conditions under which $M^{\ddagger}$ is the pseudo-inverse of M were given in [3,18] (see also [34]). The expression (6) for $M^{\ddagger}$ is a generalization to the pseudo-inverse of the block bordering method for computing the inverse of a bordered matrix [30].

If $p \geqslant q$ and if A has rank q, if $r \geqslant s$ and if D has rank s, then $S^{\dagger}S = I_s$ and it holds $M^{\ddagger}M = I_{q+s}$. So, in this case, $M^{\ddagger}$ is a left inverse of M.

If $p \leqslant q$, if A has rank p, if $r \leqslant s$ and if D has rank r, then $SS^{\dagger} = I_r$ and we have $MM^{\ddagger} = I_{p+r}$. In this case, $M^{\ddagger}$ is a right inverse of M.

We consider the system

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ d \end{pmatrix},$$

and we assume that we are in the first particular case above. Then, multiplying on the left by $M^{\ddagger}$, we obtain

$$x = A^{\dagger}b - A^{\dagger}BS^{\dagger}(d - CA^{\dagger}b), \tag{7}$$

$$y = S^{\dagger}(d - CA^{\dagger}b). \tag{8}$$

3. The new transformations

In the case of a scalar sequence, a transformation with the kernel (1) can be derived by three different approaches. First, it can be obtained by solving the system of linear equations which involves the unknown scalars. It is the *direct approach*. Then, a recursive algorithm for the implementation of the transformation could be obtained from an *elimination approach*. It consists in finding the expression of the unknowns and, then, in eliminating them from the expression of the kernel. The third approach, called the *annihilation operator approach*, isolates each unknown which is then removed by applying the forward difference operator Δ. These three approaches lead to the same transformation [11].

In this section, we will study vector sequence transformations corresponding to two different kernels. The first one will be called the *unknown vector case* since it involves unknown vectors. The second one is the *unknown matrix case* which corresponds to unknown matrices. The same types of approaches as in the scalar case will be followed. However, as we will see, they do not always lead to the same transformation.

It must be noticed that the new transformations derived below remain valid when the vectors $x, x_n \in \mathbb{R}^p$ are replaced by $p \times s$ matrices, the vectors $a_i \in \mathbb{R}^{q_i}$ in (9) by $q_i \times s$ matrices, and the vectors $z_n^i \in \mathbb{R}^{q_i}$ in (34) by $q_i \times s$ matrices. Indeed, this case corresponds to applying the transformations successively to the columns of the matrices x_n.

The role played by determinants in the solution of systems of linear equations is well known. Designants, introduced by Heyting in 1927 [29], are their counterpart for systems of linear equations in a non-commutative algebra and they are related to Schur complements. This notion is fundamental in the development of vector sequence transformations and it was extensively studied and used by Salam [37–39]. It could also be used for deriving the new transformations given in this section, as done in [15].

The forward difference operator Δ used in the sequel always acts on n which can be either a lower or an upper index.

3.1. Unknown vector case

In the *unknown vector case*, we consider a kernel with sequences of the form

$$x_n = x + Z_n^1 a_1 + \cdots + Z_n^k a_k, \quad n = 0, 1, \ldots, \tag{9}$$

where the (Z_n^i)s are known sequences of matrices in $\mathbb{R}^{p \times q_i}$ and the a_is are unknown vectors in $\mathbb{R}^{q_i}$, for $i = 1, \ldots, k$. When $\forall i$, $q_i = 1$, the unknown scalar case is recovered. If $q_i > 1$ and if we designate the columns of Z_n^i by z_n^{ij} and the components of a_i by a_{ij} for $j = 1, \ldots, q_i$, then (9) writes

$$x_n = x + a_{11} z_n^{11} + \cdots + a_{1q_1} z_n^{1q_1} + \cdots + a_{k1} z_n^{k1} + \cdots + a_{kq_k} z_n^{kq_k},$$

which is exactly the form of the vectors considered in (1). However, instead of considering each term in this expression separately, we will now treat them by block.

If we set $Z_n = [Z_n^1, \ldots, Z_n^k]$ and $a = (a_1^{\mathrm{T}}, \ldots, a_k^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{p_k}$ with $p_k = q_1 + \cdots + q_k$, then (9) takes the compact form

$$x_n = x + Z_n a.$$

If we are able to compute the vectors a_i, then, by construction, the kernel of the transformation

$$(x_n) \mapsto (y_n = x_n - Z_n^1 a_1 - \cdots - Z_n^k a_k) \tag{10}$$

will contain all sequences of the form (9).

We will now try to construct sequence transformations whose kernels consist of sequences of the form (9). Since we will obtain several transformations, the sequences they produce will be denoted by $({}_i E_{k,m_k}^{(n)})$, where k and n are as in (10), m_k is a parameter that will be introduced below, and the subscript i distinguishes the various transformations. When $m_k = k$, the subscript m_k will be suppressed.

Three main classes of approaches will be considered as in the scalar case [13].

3.1.1. *The direct approaches*

Several direct approaches will be used. As will be seen, they lead to transformations which, by construction, have all the same kernel (9). All of them can also be expressed as pseudo-Schur complements. When the matrices involved in these transformations are square and nonsingular, the pseudo-inverses are the usual inverses, and the vectors they produce can be expressed as a ratio of determinants following (4).

However, these transformations are not identical since, when applied to a sequence (x_n) different from (9), they produce distinct sequences (y_n).

First approach

Let $m_k \geqslant 1$ be an integer. In the first direct approach, we define ${}_1E_{k,m_k}^{(n)} \in \mathbb{R}^p$ as the vector satisfying the interpolation conditions

$$x_{n+i} = {}_1E_{k,m_k}^{(n)} + Z_{n+i}^1 a_1 + \cdots + Z_{n+i}^k a_k, \quad i = 0, \ldots, m_k. \tag{11}$$

The system (11) can be written as

$$\begin{pmatrix} I_p & Z_n^1 & \cdots & Z_n^k \\ I_p & Z_{n+1}^1 & \cdots & Z_{n+1}^k \\ \vdots & \vdots & & \vdots \\ I_p & Z_{n+m_k}^1 & \cdots & Z_{n+m_k}^k \end{pmatrix} \begin{pmatrix} {}_1E_{k,m_k}^{(n)} \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} x_n \\ x_{n+1} \\ \vdots \\ x_{n+m_k} \end{pmatrix}, \tag{12}$$

where the a_i depend on k and n, if (x_n) has a form which differs from (9). This system generalizes the system corresponding to the scalar case [13].

We see that we have k unknown vectors $a_i \in \mathbb{R}^{q_i}$ and the unknown vector ${}_1E_{k,m_k}^{(n)} \in \mathbb{R}^p$. Each of the $m_k + 1$ interpolation conditions (11) corresponds to p equations. In order to obtain $a_1, \ldots, a_k$ and ${}_1E_{k,m_k}^{(n)}$, we need to have, at least, $p + p_k$ equations with $p_k = q_1 + \cdots + q_k$. So, we assume that $m_k p \geqslant p_k$ (which does not mean that each q_i must be smaller than p).

Assuming that the matrix of the system (12) has rank $p + p_k$ and solving it in the least squares sense, gives the vector ${}_1E_{k,m_k}^{(n)}$

$$ {}_1E_{k,m_k}^{(n)} = [I_p, 0_{p\times q_1}, \ldots, 0_{p\times q_k}] \begin{pmatrix} I_p & Z_n^1 & \cdots & Z_n^k \\ I_p & Z_{n+1}^1 & \cdots & Z_{n+1}^k \\ \vdots & \vdots & & \vdots \\ I_p & Z_{n+m_k}^1 & \cdots & Z_{n+m_k}^k \end{pmatrix}^{\dagger} \begin{pmatrix} x_n \\ x_{n+1} \\ \vdots \\ x_{n+m_k} \end{pmatrix}, $$

where the lower indexes in I and 0 denote the dimensions (a single index is used for square matrices).

Second approach

Let us denote the matrix of the system (12) by M and partition it as in (3) with $A = I_p$ in its upper left corner. Let the right hand side be partitioned accordingly with $b = x_n$. From Formula (6), $M^{\ddagger}$ is, in general, different from $M^{\dagger}$ and it is only a left inverse of M. Multiplying (12) on the left by $M^{\ddagger}$ and using (7), we obtain a new transformation given by

$$ {}_2E_{k,m_k}^{(n)} = x_n - [Z_n^1, \ldots, Z_n^k] \begin{pmatrix} Z_{n+1}^1 - Z_n^1 & \cdots & Z_{n+1}^k - Z_n^k \\ \vdots & & \vdots \\ Z_{n+m_k}^1 - Z_n^1 & \cdots & Z_{n+m_k}^k - Z_n^k \end{pmatrix}^{\dagger} \times \begin{pmatrix} x_{n+1} - x_n \\ \vdots \\ x_{n+m_k} - x_n \end{pmatrix}. \tag{13} $$

This transformation is quite similar to the transformation defined by (15) but differs from it.

Third approach

The third direct approach consists in applying the operator Δ to (9). We obtain

$$ \Delta x_{n+i} = \Delta Z_{n+i}^1 a_1 + \cdots + \Delta Z_{n+i}^k a_k, \quad i = 0, \ldots, m_k - 1. \tag{14} $$

We now have p_k unknowns. We again assume that $m_k p \geqslant p_k$ and that the matrix of the corresponding system has rank p_k. Solving it in the least squares sense for the unknown vectors a_i and using (10), leads to the transformation

$$
{}_3E^{(n)}_{k,m_k} = x_n - [Z^1_n, \ldots, Z^k_n] \begin{pmatrix} \Delta Z^1_n & \cdots & \Delta Z^k_n \\ \vdots & & \vdots \\ \Delta Z^1_{n+m_k-1} & \cdots & \Delta Z^k_{n+m_k-1} \end{pmatrix}^{\dagger} \times \begin{pmatrix} \Delta x_n \\ \vdots \\ \Delta x_{n+m_k-1} \end{pmatrix}. \tag{15}
$$

This transformation generalizes the least squares vector sequence transformation studied in [9].

If $m_k p = p_k$, the pseudo-inverse in (15) is replaced by the inverse and, using the vector Schur determinantal formula, we obtain

$$
{}_3E^{(n)}_{k,m_k} = \frac{\begin{vmatrix} x_n & Z^1_n & \cdots & Z^k_n \\ \Delta x_n & \Delta Z^1_n & \cdots & \Delta Z^k_n \\ \vdots & \vdots & & \vdots \\ \Delta x_{n+m_k-1} & \Delta Z^1_{n+m_k-1} & \cdots & \Delta Z^k_{n+m_k-1} \end{vmatrix}}{\begin{vmatrix} \Delta Z^1_n & \cdots & \Delta Z^k_n \\ \vdots & & \vdots \\ \Delta Z^1_{n+m_k-1} & \cdots & \Delta Z^k_{n+m_k-1} \end{vmatrix}}. \tag{16}
$$

The determinant in the numerator of ${}_3E^{(n)}_{k,m_k}$ denotes the vector which is the combination of the vectors forming its first vector row (that is x_n and the columns of the matrices $Z^1_n, \ldots, Z^k_n$) obtained by the classical rule for expanding a determinant.

Third approach again

If, in (12), each block row, from the second one, is replaced by its difference with the previous one, we obtain a new transformation defined by the system

$$
\begin{pmatrix} I_p & Z^1_n & \cdots & Z^k_n \\ 0_p & \Delta Z^1_n & \cdots & \Delta Z^k_n \\ \vdots & \vdots & & \vdots \\ 0_p & \Delta Z^1_{n+m_k-1} & \cdots & \Delta Z^k_{n+m_k-1} \end{pmatrix} \begin{pmatrix} {}_3E^{(n)}_{k,m_k} \\ a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} x_n \\ \Delta x_n \\ \vdots \\ \Delta x_{n+m_k-1} \end{pmatrix}. \tag{17}
$$

Since this system splits into a system giving the a_is and an equation for ${}_3E^{(n)}_{k,m_k}$, the vectors obtained by this transformation are the same as the vectors given by (15) (this is why they have the same denomination), but they differ from those defined by (12) although the kernel of the transformation (17) is still given by (9). This is due to the fact that, when solving a system in the least squares sense, replacing an equation by its difference with another one produces a different solution.

Let us denote the matrix of the system (17) by M and partition it as in (3) with $A = I_p$ in its upper left corner and let the right hand side be partitioned accordingly with $b = x_n$. Since $C = 0_{m_k p\times p}$, we have $S = D$ and Formula (6) gives

$$M^{\ddagger} = \begin{pmatrix} I_p & -BD^{\dagger} \\ 0_{m_k p\times p} & D^{\dagger} \end{pmatrix}.$$

So $M^{\ddagger} = M^{\dagger}$, and (15) is recovered from (7).

Fourth approach

Let us now discuss another direct approach which follows more closely the formulation of the vector E-transformation [5].

Let $Y \in \mathbb{R}^{p\times m}$. We consider the $m_k \times m_k$ block diagonal matrix $\widetilde{Y}$ which consists of m_k blocks Y on its diagonal. We assume that $m_k m \geqslant p_k = q_1 + \cdots + q_k$. Multiplying on the left the system (14) by $\widetilde{Y}^{\mathrm{T}}$ gives

$$\begin{pmatrix} Y^{\mathrm{T}}\Delta Z_n^1 & \cdots & Y^{\mathrm{T}}\Delta Z_n^k \\ \vdots & & \vdots \\ Y^{\mathrm{T}}\Delta Z_{n+m_k-1}^1 & \cdots & Y^{\mathrm{T}}\Delta Z_{n+m_k-1}^k \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} = \begin{pmatrix} Y^{\mathrm{T}}\Delta x_n \\ \vdots \\ Y^{\mathrm{T}}\Delta x_{n+m_k-1} \end{pmatrix}.$$

Solving this system in the least squares sense, leads to the transformation

$$\begin{aligned} {}_4E_{k,m_k}^{(n)} = x_n - [Z_n^1, \ldots, Z_n^k] &\begin{pmatrix} Y^{\mathrm{T}}\Delta Z_n^1 & \cdots & Y^{\mathrm{T}}\Delta Z_n^k \\ \vdots & & \vdots \\ Y^{\mathrm{T}}\Delta Z_{n+m_k-1}^1 & \cdots & Y^{\mathrm{T}}\Delta Z_{n+m_k-1}^k \end{pmatrix}^{\dagger} \\ &\times \begin{pmatrix} Y^{\mathrm{T}}\Delta x_n \\ \vdots \\ Y^{\mathrm{T}}\Delta x_{n+m_k-1} \end{pmatrix}. \end{aligned} \tag{18}$$

The matrix Y can depend on k and n, thus leading to another transformation with the same kernel.

If $\forall i$, $q_i = 1$, if $m_k = k$, and if $m = 1$, the vector E-transformation is recovered. It can be recursively implemented by the vector E-algorithm [5].

Fifth approach

In the fourth approach, m_k equations of the form (14) were used and the transformation (18) was defined from a single auxiliary matrix Y. Another approach consists in using one single relation of the form (9) and several auxiliary matrices Y_i. Such an approach follows the lines leading to the RPA and its variants [6].

Let $Y_i \in \mathbb{R}^{p\times r_i}$ and $\widehat{Y} = [Y_1, \ldots, Y_{m_k}]$. We consider the relation

$$\Delta Z_n^1 a_1 + \cdots + \Delta Z_n^k a_k = \Delta x_n,$$

and we multiply it by $\widehat{Y}^{\mathrm{T}}$. We obtain a system of $r_1 + \cdots + r_{m_k}$ equations in $p_k = q_1 + \cdots + q_k$ unknowns. Assuming that $r_1 + \cdots + r_{m_k} \geqslant p_k$ and $p \geqslant p_k$, and solving the system in the least squares sense leads to the transformation

$$
{}_5E_{k,m_k}^{(n)} = x_n - [Z_n^1, \ldots, Z_n^k] \begin{pmatrix} Y_1^{\mathrm{T}} \Delta Z_n^1 & \cdots & Y_1^{\mathrm{T}} \Delta Z_n^k \\ \vdots & & \vdots \\ Y_{m_k}^{\mathrm{T}} \Delta Z_n^1 & \cdots & Y_{m_k}^{\mathrm{T}} \Delta Z_n^k \end{pmatrix}^{\dagger} \begin{pmatrix} Y_1^{\mathrm{T}} \Delta x_n \\ \vdots \\ Y_{m_k}^{\mathrm{T}} \Delta x_n \end{pmatrix}. \tag{19}
$$

The matrices Y_i can depend on k and n, thus leading to another transformation with the same kernel. We see that, due to the condition $p \geqslant p_k = q_1 + \cdots + q_k$, the square case, $\forall i$, $q_i = p$, cannot be handled by this approach.

If $\forall i$, $q_i = r_i = 1$ and $m_k = k$, the vector transformations given in [6,14,31] are recovered. They can be implemented by recursive algorithms.

Other approaches and discussion

Transformations similar to (18) and (19) can be obtained from (11) instead of (14) and using matrices in the style of $\widetilde{Y}$ and $\widehat{Y}$ respectively. The corresponding systems can also be solved by multiplying them on the left by $M^{\ddagger}$, thus leading to transformations in the style of (13). In addition, transformations can be constructed by some mixed strategies between these last two approaches as in [8].

By construction, the kernels of all the transformations defined in this section, for k fixed, contain the sequences of the form (9). In fact, more complete results can be proved. Let us consider the first direct approach and let $G_{k,i}^{(n)}$ be the solution of the system

$$
\begin{pmatrix} I_p & Z_n^1 & \cdots & Z_n^k \\ I_p & Z_{n+1}^1 & \cdots & Z_{n+1}^k \\ \vdots & \vdots & & \vdots \\ I_p & Z_{n+m_k}^1 & \cdots & Z_{n+m_k}^k \end{pmatrix} \begin{pmatrix} G_{k,i}^{(n)} \\ b_1 \\ \vdots \\ b_k \end{pmatrix} = \begin{pmatrix} Z_n^i \\ Z_{n+1}^i \\ \vdots \\ Z_{n+m_k}^i \end{pmatrix}, \tag{20}
$$

where the b_i depend on k, n and m_k. This system is the same as the system (12) after replacing, in its right hand side, x_n by Z_n^i. The same change can be made in the systems corresponding to the other direct approaches for defining the matrices $G_{k,i}^{(n)}$ in a similar way. Thus, the following result holds:

Theorem 1. *If*

$$x_n = x + Z_n^1 a_1 + Z_n^2 a_2 + \cdots, \quad n = 0, 1, \ldots,$$

with $a_j \in \mathbb{R}^{q_j}$ *then, for* $k = 0, 1, \ldots$ *and* $m_k = k$, *it holds, for the transformations* $i = 1, \ldots, 5$ (*assuming that the conditions on the dimensions and ranks of the matrices involved are satisfied*),

$${}_iE_k^{(n)} = x + G_{k,k+1}^{(n)} a_{k+1} + G_{k,k+2}^{(n)} a_{k+2} + \cdots, \quad n = 0, 1, \ldots,$$

where the matrices $G^{(n)}_{k,k+1}, G^{(n)}_{k,k+2}, \ldots$ *are given by the same system as the system for* ${}_iE^{(n)}_k$ *after replacing, in its right hand side,* x_n *by* Z^i_n.

Proof. Let us consider the transformation defined by ${}_1E^{(n)}_k$. We only have to separate the terms in x_n as

$$x_n = (x + Z^1_n a_1 + \cdots + Z^k_n a_k) + (Z^{k+1}_n a_{k+1} + \cdots),$$

to replace x_n in the right hand side of (12) by this expression, and to make use of (20).

The proofs for the other direct approaches are similar. □

Although the transformations obtained by the preceding direct approaches have the same kernel, they do not coincide because each of them comes out from the pseudo-inverse of a different matrix.

Recursive algorithms for the implementation of the transformations defined in this section remain to be obtained. It must be noticed that similar vector sequence transformations can be defined even when the number of equations is smaller than the number of unknowns (that is if $m_k p < p_k$). However, in this case, their kernels are no longer given by (9).

3.1.2. The elimination approaches

In this section, we assume that $\forall i,\ p \geqslant q_i$ and that all $p \times q_i$ matrices whose pseudo-inverse is needed have rank q_i. Notice that summing up these inequalities leads to $m_k p \geqslant q_1 + \cdots + q_k = p_k$, a condition already encountered in Section 3.1.1.

As it will be easy to check, the new transformations that will be obtained now use the same number of vectors x_n and matrices Z^i_n as the transformations of the previous section when $m_k = k$. So, the new transformations will no longer depend on an integer m_k and this subscript will be suppressed.

We now assume that

$$x_n = x + Z^1_n a_1 + Z^2_n a_2 + Z^3_n a_3 + \cdots, \quad n = 0, 1, \ldots, \tag{21}$$

with possibly an infinite number of terms in this expression.

Several elimination approaches can be studied.

First approach

Let us write down (21) for the index $n + 1$, multiply it by $Z^1_n (Z^1_{n+1})^\dagger$ and subtract it from x_n. We get

$$\begin{aligned} x_n - Z^1_n (Z^1_{n+1})^\dagger x_{n+1} = {} & \big(I_p - Z^1_n (Z^1_{n+1})^\dagger\big)x + \big(Z^1_n - Z^1_n (Z^1_{n+1})^\dagger Z^1_{n+1}\big)a_1 \\ & + \big(Z^2_n - Z^1_n (Z^1_{n+1})^\dagger Z^2_{n+1}\big)a_2 + \cdots \end{aligned}$$

Since $(Z_{n+1}^1)^\dagger Z_{n+1}^1 = I_{q_1}$, the term in a_1 disappears. Then, we multiply both sides of this relation by the inverse of the matrix $\left(I_p - Z_n^1(Z_{n+1}^1)^\dagger\right)$ and we set

$$\begin{aligned}
{}_6E_1^{(n)} &= \left[I_p - Z_n^1(Z_{n+1}^1)^\dagger\right]^{-1}\left(x_n - Z_n^1(Z_{n+1}^1)^\dagger x_{n+1}\right),\\
G_{1,i}^{(n)} &= \left[I_p - Z_n^1(Z_{n+1}^1)^\dagger\right]^{-1}\left(Z_n^i - Z_n^1(Z_{n+1}^1)^\dagger Z_{n+1}^i\right), \quad i > 1.
\end{aligned}$$

We get

$${}_6E_1^{(n)} = x + G_{1,2}^{(n)}a_2 + G_{1,3}^{(n)}a_3 + \cdots$$

Obviously, the same process can be repeated. We set ${}_6E_0^{(n)} = x_n$ and $G_{0,i}^{(n)} = Z_n^i$ and we assume that, for some $k \geqslant 1$, we have

$${}_6E_{k-1}^{(n)} = x + G_{k-1,k}^{(n)}a_k + G_{k-1,k+1}^{(n)}a_{k+1} + \cdots$$

For $k = 1$, this relation reduces to (21). We multiply ${}_6E_{k-1}^{(n+1)}$ on the left by $G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger$ and we subtract it from ${}_6E_{k-1}^{(n)}$. The term $G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger a_k$ disappears and we get

$$\begin{aligned}
{}_6E_{k-1}^{(n)} - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger {}_6E_{k-1}^{(n+1)} &= \left(I_p - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger\right)x\\
&+ \left(G_{k-1,k+1}^{(n)} - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger G_{k-1,k+1}^{(n+1)}\right)a_{k+1} + \cdots
\end{aligned}$$

We multiply both sides by the inverse of $\left(I_p - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger\right)$ and we set

$${}_6E_k^{(n)} = \left[I_p - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger\right]^{-1}\left[{}_6E_{k-1}^{(n)} - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger {}_6E_{k-1}^{(n+1)}\right] \tag{22}$$

$$G_{k,i}^{(n)} = \left[I_p - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger\right]^{-1}\left[G_{k-1,i}^{(n)} - G_{k-1,k}^{(n)}\left(G_{k-1,k}^{(n+1)}\right)^\dagger G_{k-1,i}^{(n+1)}\right], \quad i > k. \tag{23}$$

So, it holds

$${}_6E_k^{(n)} = x + G_{k,k+1}^{(n)}a_{k+1} + G_{k,k+2}^{(n)}a_{k+2} + \cdots$$

Thus, by construction, the kernel of this transformation contains the sequences of the form (9), since for such a sequence $a_i = 0$ for $i > k$. Also notice that ${}_6E_k^{(n)}$ can be expressed as a pseudo-Schur complement and that we have $G_{k,i}^{(n)} = 0$ for $i \leqslant k$.

Second approach

Let us now consider another elimination approach. Since it is quite similar to the previous approach, it will be described with less details.

We set ${}_7E_0^{(n)} = x_n$ and $G_{0,i}^{(n)} = Z_n^i$. So, (21) can be written as

$$ {}_7E_{k-1}^{(n)} = x + G_{k-1,k}^{(n)} a_k + G_{k-1,k+1}^{(n)} a_{k+1} + \cdots \tag{24} $$

for $k = 1$. Let us assume that this relation holds for some k. We apply the operator Δ to (24) and we multiply on the left by the pseudo-inverse of $\Delta G_{k-1,k}^{(n)}$. Since $\left(\Delta G_{k-1,k}^{(n)}\right)^{\dagger} \Delta G_{k-1,k}^{(n)} = I_{q_k}$, we obtain the expression of a_k

$$ \begin{aligned} a_k = {} & \left(\Delta G_{k-1,k}^{(n)}\right)^{\dagger} \Delta {}_7E_{k-1}^{(n)} - \left(\Delta G_{k-1,k}^{(n)}\right)^{\dagger} \Delta G_{k-1,k+1}^{(n)} a_{k+1} \\ & - \left(\Delta G_{k-1,k}^{(n)}\right)^{\dagger} \Delta G_{k-1,k+2}^{(n)} a_{k+2} - \cdots \end{aligned} $$

Let us define ${}_7E_k^{(n)}$ and $G_{k,i}^{(n)}$ by (notice that the matrices $G_{k,i}^{(n)}$ are different from those of the first elimination approach)

$$ {}_7E_k^{(n)} = {}_7E_{k-1}^{(n)} - G_{k-1,k}^{(n)} \left(\Delta G_{k-1,k}^{(n)}\right)^{\dagger} \Delta {}_7E_{k-1}^{(n)}, \tag{25} $$

$$ G_{k,i}^{(n)} = G_{k-1,i}^{(n)} - G_{k-1,k}^{(n)} \left(\Delta G_{k-1,k}^{(n)}\right)^{\dagger} \Delta G_{k-1,i}^{(n)}, \quad i > k. \tag{26} $$

Replacing a_k by its expression in (24) gives

$$ {}_7E_k^{(n)} = x + G_{k,k+1}^{(n)} a_{k+1} + G_{k,k+2}^{(n)} a_{k+2} + \cdots $$

We can check that ${}_7E_k^{(n)} \in \mathbb{R}^p$ and $G_{k,i}^{(n)} \in \mathbb{R}^{p \times q_i}$. By induction, we see that $\forall n$, $G_{k,i}^{(n)} = 0$ for $i < k$. Moreover, from (26), we have $\forall n$, $G_{k,k}^{(n)} = 0$.

The relations (25) and (26) are similar to the rules of the E-algorithm [5]. Due to the non-commutativity of the matrix products, it does not seem that the preceding rules could be simplified. A similar situation arose in the general matrix case studied in [15, p. 507]. We also see that ${}_7E_k^{(n)}$ and $G_{k,i}^{(n)}$ respectively appear as the pseudo-Schur complements of $\Delta G_{k-1,k}^{(n)}$ in

$$ \begin{pmatrix} {}_7E_{k-1}^{(n)} & G_{k-1,k}^{(n)} \\ \Delta {}_7E_{k-1}^{(n)} & \Delta G_{k-1,k}^{(n)} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} G_{k-1,i}^{(n)} & G_{k-1,k}^{(n)} \\ \Delta G_{k-1,i}^{(n)} & \Delta G_{k-1,k}^{(n)} \end{pmatrix}. $$

By construction, as in the first elimination approach, we see that all sequences of the form (9) belong to the kernel of this transformation.

Other approaches

The second elimination approach can be generalized as follows. Let $U_n^k \in \mathbb{R}^{p \times r_k}$ with $r_k \geqslant q_k$ be matrices which can depend on n and such that $(U_n^k)^{\mathrm{T}} \Delta Z_n^k$ has rank q_k. Applying the operator Δ to (21) and multiplying by $(U_n^1)^{\mathrm{T}}$, we obtain

$$ (U_n^1)^{\mathrm{T}} \Delta x_n = (U_n^1)^{\mathrm{T}} \Delta Z_n^1 a_1 + (U_n^1)^{\mathrm{T}} \Delta Z_n^2 a_2 + (U_n^1)^{\mathrm{T}} \Delta Z_n^3 a_3 + \cdots $$

We obtain

$$a_1 = \left[(U_n^1)^{\mathrm{T}} \Delta Z_n^1\right]^{\dagger} (U_n^1)^{\mathrm{T}} \Delta x_n - \left[(U_n^1)^{\mathrm{T}} \Delta Z_n^1\right]^{\dagger} (U_n^1)^{\mathrm{T}} \Delta Z_n^2 a_2 \\ - \left[(U_n^1)^{\mathrm{T}} \Delta Z_n^1\right]^{\dagger} (U_n^1)^{\mathrm{T}} \Delta Z_n^3 a_3 - \cdots$$

Thus

$$x_n = x + Z_n^1\left[(U_n^1)^{\mathrm{T}} \Delta Z_n^1\right]^{\dagger} (U_n^1)^{\mathrm{T}} \Delta x_n - Z_n^1\left[(U_n^1)^{\mathrm{T}} \Delta Z_n^1\right]^{\dagger} (U_n^1)^{\mathrm{T}} \Delta Z_n^2 a_2 \\ - Z_n^1\left[(U_n^1)^{\mathrm{T}} \Delta Z_n^1\right]^{\dagger} (U_n^1)^{\mathrm{T}} \Delta Z_n^3 a_3 - \cdots + Z_n^2 a_2 + Z_n^3 a_3 + \cdots$$

We set $E_0(v_n) = v_n$ and

$$E_k(v_n) = E_{k-1}(v_n) - E_{k-1}(Z_n^k)\left[(U_n^k)^{\mathrm{T}} \Delta E_{k-1}(Z_n^k)\right]^{\dagger} (U_n^k)^{\mathrm{T}} \Delta E_{k-1}(v_n), \tag{27}$$

where $v_n \in \mathbb{R}^{p\times r}$ with $r \geqslant 1$. Thus, the preceding relation writes

$$E_k(x_n) = x + E_k(Z_n^{k+1})a_{k+1} + E_k(Z_n^{k+2})a_{k+2} + \cdots \tag{28}$$

for $k = 1$. As for the other transformations, we set ${}_8E_k^{(n)} = E_k(x_n)$ and $G_{k,i}^{(n)} = E_k(Z_n^i)$, $i > k$, and it can be proved by induction that (28) holds for all k. This approach leads to a transformation with kernel (9) only in the case where the matrices Z_n^i are square, that is $\forall i, q_i = p$.

The matrix $\left[(U_n^k)^{\mathrm{T}} \Delta E_{k-1}(Z_n^k)\right]^{\dagger} (U_n^k)^{\mathrm{T}}$ appears as a pseudo left inverse of $\Delta E_{k-1}(Z_n^k)$. We see that the algorithm (25)–(26) is recovered for the choice $U_n^k = \Delta E_{k-1}(Z_n^k)$. Thus, the multilevel VTT approach introduced in [12] has been generalized. Moreover, if the matrices U_n^k are such that $(U_n^k)^{\mathrm{T}} E_{k-1}(Z_{n+1}^k) = 0$, the preceding formulae simplify and we obtain a generalization of the multilevel biorthogonal VTT (BVTT) given in [12].

Obviously, the first elimination approach could be generalized similarly. Its kernel will also contain all sequences given by (9) and the same results hold.

For all the transformations obtained by the elimination approach, we have, by construction, the following result:

Theorem 2. *If*

$$x_n = x + Z_n^1 a_1 + Z_n^2 a_2 + \cdots, \quad n = 0, 1, \ldots,$$

with $a_j \in \mathbb{R}^{q_j}$, *then, for* $k = 0, 1, \ldots,$ *it holds, for the transformations* $i = 6, 7$ *and* 8 (*assuming that the conditions on the dimensions and ranks of the matrices involved are satisfied*),

$${}_iE_k^{(n)} = x + G_{k,k+1}^{(n)} a_{k+1} + G_{k,k+2}^{(n)} a_{k+2} + \cdots, \quad n = 0, 1, \ldots,$$

where the matrices $G_{k,k+1}^{(n)}, G_{k,k+2}^{(n)}, \ldots$ *are given by* (23) *for* $i = 6$, *by* (26) *for* $i = 7$, *and by* (27) *for* $i = 8$, *respectively.*

3.1.3. The annihilation operator approach

We will now try to follow exactly the annihilation operator approach explained in [11] for the scalar case. We will see that this approach can also be simplified.

Let (v_n) be a sequence of matrices in $\mathbb{R}^{p\times r}$ where $r \geqslant 1$. We define the linear difference operators N_k acting on the sequence (v_n) by, $\forall n$,

$$N_0(v_n) = v_n,$$
$$N_{k+1}(v_n) = \Delta\left(\left[N_k(Z_n^{k+1})\right]^\dagger N_k(v_n)\right).$$

Let us assume that

$$x_n = x + Z_n^1 a_1 + Z_n^2 a_2 + \cdots, \quad n = 0, 1, \ldots, \tag{29}$$

which can be written as

$$N_0(x_n) = N_0(I)x + N_0(Z_n^1)a_1 + N_0(Z_n^2)a_2 + \cdots,$$

where I is the identity matrix of dimension p.

Multiplying both sides on the left by $[N_0(Z_n^1)]^\dagger$ gives

$$[N_0(Z_n^1)]^\dagger N_0(x_n) = [N_0(Z_n^1)]^\dagger N_0(I)x + a_1 + [N_0(Z_n^1)]^\dagger N_0(Z_n^2)a_2 + \cdots$$

Applying the operator Δ and since $\Delta a_1 = 0$, we obtain

$$\Delta\left([N_0(Z_n^1)]^\dagger N_0(x_n)\right) = \Delta\left([N_0(Z_n^1)]^\dagger N_0(I)\right)x + \Delta\left([N_0(Z_n^1)]^\dagger N_0(Z_n^2)\right)a_2 + \cdots$$

that is

$$N_1(x_n) = N_1(I)x + N_1(Z_n^2)a_2 + N_1(Z_n^3)a_3 + \cdots$$

It is easy to check that $N_1(x_n) \in \mathbb{R}^{q_1}$, $N_1(I) \in \mathbb{R}^{q_1\times p}$ and $N_1(Z_n^i) \in \mathbb{R}^{q_1\times q_i}$ for $i \geqslant 2$.

Obviously, the same process can be repeated under the assumptions that $q_{k+1} \leqslant q_k$ (a new condition that has to be introduced) and that $N_k(Z_n^{k+1})$ has rank q_{k+1} (in order that $(N_k(Z_n^{k+1}))^\dagger N_k(Z_n^{k+1}) = I$), and it leads to

$$N_k(x_n) = N_k(I)x + N_k(Z_n^{k+1})a_{k+1} + N_k(Z_n^{k+2})a_{k+2} + \cdots \tag{30}$$

with $N_k(x_n) \in \mathbb{R}^{q_k}$, $N_k(I) \in \mathbb{R}^{q_k\times p}$ and $N_k(Z_n^i) \in \mathbb{R}^{q_k\times q_i}$ for $i \geqslant k+1$.

We consider the sequence transformation

$$E_k : (v_n) \mapsto (E_k(v_n))$$

defined by

$$E_k(v_n) = [N_k(I)]^\dagger N_k(v_n), \quad n = 0, 1, \ldots$$

Multiplying both sides of (30) by $[N_k(I)]^\dagger$, we obtain

$$E_k(x_n) = E_k(I)x + E_k(Z_n^{k+1})a_{k+1} + E_k(Z_n^{k+2})a_{k+2} + \cdots \tag{31}$$

It can be checked that $E_k(x_n) \in \mathbb{R}^p$ and $E_k(Z_n^i) \in \mathbb{R}^{p\times q_i}$ for $i \geqslant k+1$.

Let us now show how to compute recursively the vectors $E_k(x_n)$. We have

$$\begin{aligned} E_k(v_n) &= [N_k(I)]^\dagger N_k(v_n) \\ &= \left[\Delta\left([N_{k-1}(Z_n^k)]^\dagger N_{k-1}(I)\right)\right]^\dagger \Delta\left([N_{k-1}(Z_n^k)]^\dagger N_{k-1}(v_n)\right). \end{aligned}$$

Since $N_{k-1}(I)N_{k-1}(I)^\dagger = I$, the second relation can also be written as

$$\begin{aligned} E_k(v_n) = {} & \left[\Delta\left([N_{k-1}(Z_n^k)]^\dagger N_{k-1}(I)\right)\right]^\dagger \\ & \times \Delta\left([N_{k-1}(Z_n^k)]^\dagger N_{k-1}(I)N_{k-1}(I)^\dagger N_{k-1}(v_n)\right) \end{aligned}$$

that is, if $N_{k-1}(Z_n^k)$ has rank q_{k-1}, and since, with our assumptions, $[N_{k-1}(I)^\dagger N_{k-1}(Z_n^k)]^\dagger = [N_{k-1}(Z_n^k)]^\dagger N_{k-1}(I)$, we obtain

$$\begin{aligned} E_k(v_n) = {} & \left[E_{k-1}(Z_{n+1}^k)^\dagger - E_{k-1}(Z_n^k)^\dagger\right]^\dagger \\ & \times \left[E_{k-1}(Z_{n+1}^k)^\dagger E_{k-1}(v_{n+1}) - E_{k-1}(Z_n^k)^\dagger E_{k-1}(v_n)\right]. \end{aligned} \tag{32}$$

The relation (32) generalizes the rule of the E-algorithm [5]. The vectors ${}_9E_k^{(n)}$ are computed by replacing v_n by x_n in this rule. We also need to compute the auxiliary matrices $G_{k,i}^{(n)} = E_k(Z_n^i)$, which are obtained by setting $v_n = Z_n^i$ in (32). Due to the non-commutativity of the matrix products, it does not seem that the preceding rule could be written in a different way. Notice that a similar situation arose in the general matrix case studied in [15, p. 507].

An important point to notice is that, in this approach, we had to introduce the assumptions that, for all k, the matrix $N_k(Z_n^{k+1}) \in \mathbb{R}^{q_k\times q_{k+1}}$ has rank q_k and that $q_{k+1} \leqslant q_k$. Obviously, these assumptions can only be true if all the q_i are equal.

Notice that $N_k(I)x = N_k(x)$ and $E_k(I)x = E_k(x)$. However, in general, $(N_k(I))^\dagger N_k(I) \neq I$ and, thus, $E_k(I)$ in (31) is not the identity. So, this approach leads to an algorithm whose kernel is the set of sequences of the form (9) only if the matrices Z_n^i are square and nonsingular.

This approach can also be recovered without introducing the operators N_k. We set $E_0(v_n) = v_n$ and we assume that

$$E_{k-1}(x_n) = E_{k-1}(x) + E_{k-1}(Z_n^k)a_k + E_{k-1}(Z_n^{k+1})a_{k+1} + \cdots$$

Indeed (29) corresponds to this form for $k = 1$. Multiplying both sides by $(E_{k-1}(Z_n^k))^\dagger$, we obtain

$$\begin{aligned} (E_{k-1}(Z_n^k))^\dagger E_{k-1}(x_n) = {} & (E_{k-1}(Z_n^k))^\dagger E_{k-1}(x) + a_k \\ & + (E_{k-1}(Z_n^k))^\dagger E_{k-1}(Z_n^{k+1})a_{k+1} + \cdots \end{aligned}$$

since, with our assumptions, $(E_{k-1}(Z_n^k))^\dagger E_{k-1}(Z_n^k) = I$. In order to suppress a_k from this expression, let us apply the operator Δ to both sides. We get

$$\Delta\left[(E_{k-1}(Z_n^k))^\dagger E_{k-1}(x_n)\right] = \Delta\left[(E_{k-1}(Z_n^k))^\dagger E_{k-1}(x)\right] + \Delta\left[(E_{k-1}(Z_n^k))^\dagger E_{k-1}(Z_n^{k+1})\right]a_{k+1} + \cdots$$

Defining $E_k(v_n) \in \mathbb{R}^{p\times r}$ by

$$E_k(v_n) = \left[\Delta(E_{k-1}(Z_n^k))^\dagger\right]^\dagger \Delta\left[(E_{k-1}(Z_n^k))^\dagger E_{k-1}(v_n)\right], \tag{33}$$

the preceding relation writes

$$E_k(x_n) = E_k(x) + E_k(Z_n^{k+1})a_{k+1} + \cdots$$

We see that (33) is exactly (32) and that, for this second form of the annihilation operator approach, it was not necessary to introduce the additional restriction $q_{k+1} \leqslant q_k$.

However, as above, it must be noticed that $E_k(x) \neq x$ and so the transformation defined by ${}_9E_k^{(n)} = E_k(x_n)$ and $G_{k,i}^{(n)} = E_k(Z_n^i)$ does not have a kernel of the form (9), except in the case of square matrices Z_n^i. Obviously, in this case, the pseudo-inverses in (33) have to be replaced by the inverses (assuming that they exist). When the matrices Z_n^i are rectangular, the kernel of this transformation is unknown. Thus, we have the following result, whose proof is easy.

Theorem 3. *If*

$$x_n = x + Z_n^1 a_1 + Z_n^2 a_2 + \cdots, \quad n = 0, 1, \ldots,$$

with $a_i \in \mathbb{R}^p$, $Z_n^i \in \mathbb{R}^{p\times p}$, and if all matrices whose inverse is needed are nonsingular, then, for $k = 0, 1, \ldots,$ it holds (assuming that the conditions on the dimensions and ranks of the matrices involved are satisfied)

$${}_9E_k^{(n)} = x + G_{k,k+1}^{(n)} a_{k+1} + G_{k,k+2}^{(n)} a_{k+2} + \cdots, \quad n = 0, 1, \ldots,$$

where the matrices $G_{k,j}^{(n)} = E_k(Z_n^j)$ are given by (33).

3.1.4. Comparison of the approaches

We will now compare the three preceding classes of approaches.

We saw that, if the sequence (x_n) has the form (9), then, by construction, $\forall n, {}_iE_k^{(n)} = x$ for $i = 1, \ldots, 8$ (if the conditions on the dimensions and the ranks of the matrices involved in each transformation are satisfied). Thus, these transformations have the same kernel. However, since the intermediate vectors ${}_iE_j^{(n)}$ are not identical for $j < k$ (except for ${}_2E_1^n$, ${}_3E_1^n$ and ${}_7E_1^n$) these transformations do not coincide. Notice that the condition $p \geqslant p_k$ cannot be satisfied by the fifth direct approach (Formula (19)) in the square case $\forall i, q_i = p$.

As explained above, if (x_n) satisfies (9), $\forall n$, ${}_9E_k^{(n)} \neq x$, except in the square case. So, in the rectangular case, the transformation (32) has a different (and unknown) kernel.

Let us now compare the three types of approaches when $m_k = k$, a condition needed so that the various approaches use exactly the same number of vectors x_n and matrices Z_n^i. We also assume that $kp = p_k = q_1 + \cdots + q_k$. Since the algorithms obtained by the elimination and the annihilation approaches are recursive, this condition must be satisfied for all k, which means that $q_k = p$ for all k. So, we are in the square case. If all the q_is are equal to p, pseudo-inverses become classical inverses, the product of a matrix by its inverse is commutative, and the inverse of a product is the product of the inverses in the reverse order. It follows that, for all k and n, the vectors ${}_7E_k^{(n)}$ obtained by (25) and the vectors ${}_9E_k^{(n)}$ coming out from (32) (or (33)) are identical. The same is true for the matrices $G_{k,i}^{(n)}$. Indeed, Formula (25) has the structure

$$z = y - B(A - B)^{-1}(x - y),$$

where A and B are square nonsingular matrices, and x, y and z are vectors (or matrices). Thus, we have

$$\begin{aligned} z &= y - [(A - B)B^{-1}]^{-1}(x - y) = y - [AB^{-1} - BB^{-1}]^{-1}(x - y) \\ &= y - [A(B^{-1} - A^{-1})]^{-1}(x - y) = y - (B^{-1} - A^{-1})^{-1}A^{-1}(x - y) \\ &= (B^{-1} - A^{-1})^{-1}[(B^{-1} - A^{-1})y - A^{-1}x + A^{-1}y] \\ &= (B^{-1} - A^{-1})^{-1}[B^{-1}y - A^{-1}x], \end{aligned}$$

which has the same structure as Formula (32) (or (33)). This result can also be proved by using the identity

$$(B - A)^{-1} = B^{-1} - B^{-1}(B^{-1} - A^{-1})^{-1}B^{-1}.$$

It is also easy to see that (22) and (25) are equivalent. Moreover, in the square case, we see that the vectors ${}_iE_k(x_n)$ are identical for $i = 1, 2, 3$ and 4. They also coincide with the vectors ${}_6E_k(x_n)$, ${}_7E_k(x_n)$, and ${}_9E_k(x_n)$.

So, even if (x_n) does not have the form (9), all the results provided by the formulae (12), (13), (15) (or (17)) with $m_k = k$, (18) with $m_k = k$ and $m = p$, (22), (25), and (32) (or (33)) are identical and they define the same sequence transformation, in the case of square matrices Z_n^i. The transformations (18) with $m \neq p$, (19), and (27) are different and the vectors they compute do not coincide with those given by the other transformations.

As a particular case, let (Z_n) be a sequence of $p \times p$ nonsingular matrices and let $Z_n^i = (Z_n)^i$. Our transformations (some of which are identical) define generalizations of the well-known Richardson extrapolation process. If $\forall i$, $q_i = 1$ and if $Z_n^i = \Delta x_{n+i-1}$, our transformations generalize the Shanks transformation, that is the ε-algorithm. Moreover, if x_n is the nth partial sum of a formal power series in the complex variable z, generalizations of the Padé approximants are obtained. On these topics, see [10, pp. 78–95, 216–228] and [16,17].

3.2. Unknown matrix case

In the *unknown matrix case*, we consider a kernel consisting of sequences of the form

$$x_n = x + A_1 z_n^1 + \cdots + A_k z_n^k, \quad n = 0, 1, \ldots, \tag{34}$$

where the A_is are unknown $p \times q_i$ matrices and the (z_n^i)s are known sequences in $\mathbb{R}^{q_i}$ for $i = 1, \ldots, k$. This case is a generalization of the kernel defined by (1). It was also already considered in [15, Section 5] with all the q_is identical. The case where the matrices A_i are square and diagonal can be treated by applying the scalar E-algorithm separately to each component of the vectors x_n.

As in the unknown vector case, if we are able to compute the matrices A_i, then, by construction, the kernel of the transformation

$$(x_n) \mapsto (y_n = x_n - A_1 z_n^1 - \cdots - A_k z_n^k)$$

will contain all sequences of the form (34).

As in the unknown vector case, we will now try to construct, by three different approaches, transformations whose kernels consist of sequences of the form (34). The sequences obtained will be denoted as in the unknown vector case.

3.2.1. The direct approach

We set

$$A = [A_1, \ldots, A_k] \in \mathbb{R}^{p \times p_k}, \quad z_n = \begin{pmatrix} z_n^1 \\ \vdots \\ z_n^k \end{pmatrix} \in \mathbb{R}^{p_k},$$

where $p_k = q_1 + \cdots + q_k$.

Thus the sequences of the kernel can be written as

$$x_n = x + A z_n, \quad n = 0, 1, \ldots, \tag{35}$$

which shows that the approach of [15] can be used. However, we will now present another way of treating this case. This approach follows the lines of the third direct approach used in the unknown vector case.

We apply the operator Δ to (35) and we consider the relations

$$\Delta x_{n+i} = A \Delta z_{n+i}, \quad i = 0, \ldots, m_k - 1,$$

with $m_k \geqslant p_k$. They can be written as

$$[\Delta x_n, \ldots, \Delta x_{n+m_k-1}] = A[\Delta z_n, \ldots, \Delta z_{n+m_k-1}].$$

Assuming that the $p_k \times m_k$ matrix $[\Delta z_n, \ldots, \Delta z_{n+m_k-1}]$ has rank p_k and solving this system in the least squares sense for the matrix A leads to the transformation

$$ {}_{10}E_{k,m_k}^{(n)} = x_n - [\Delta x_n, \ldots, \Delta x_{n+m_k-1}] \begin{pmatrix} \Delta z_n^1 & \cdots & \Delta z_{n+m_k-1}^1 \\ \vdots & & \vdots \\ \Delta z_n^k & \cdots & \Delta z_{n+m_k-1}^k \end{pmatrix}^{\dagger} \begin{pmatrix} z_n^1 \\ \vdots \\ z_n^k \end{pmatrix}. \tag{36} $$

The vector ${}_{10}E_{k,m_k}^{(n)}$ can be expressed as a pseudo-Schur complement.

Since we have pp_k unknowns, the matrices $A_1, \ldots, A_k$ can only be exactly computed if we write down, at least, the same number of equations, that is if $m_k \geqslant p_k$. By construction, the kernel of the transformation $(x_n) \mapsto ({}_{10}E_{k,m_k}^{(n)})$, for k fixed, contains the sequences of the form (34). It must be noticed that this transformation can also be defined by (36) even for $m_k < p_k$. However, in this case, its kernel is no longer given by (34).

If $m_k = p_k$, the Schur complement leads to the determinantal formula

$$ {}_{10}E_{k,p_k}^{(n)} = \begin{vmatrix} x_n & \Delta x_n & \cdots & \Delta x_{n+p_k-1} \\ z_n^1 & \Delta z_n^1 & \cdots & \Delta z_{n+p_k-1}^1 \\ \vdots & \vdots & & \vdots \\ z_n^k & \Delta z_n^k & \cdots & \Delta z_{n+p_k-1}^k \end{vmatrix} \Bigg/ \begin{vmatrix} \Delta z_n^1 & \cdots & \Delta z_{n+p_k-1}^1 \\ \vdots & & \vdots \\ \Delta z_n^k & \cdots & \Delta z_{n+p_k-1}^k \end{vmatrix}. $$

This formula, although it has a form quite similar to (16), means that ${}_{10}E_{k,p_k}^{(n)}$ is a linear combination of the vectors $x_n, \Delta x_n, \ldots, \Delta x_{n+p_k-1}$ which form its first vector row.

In the unknown matrix case, other direct approaches following the lines of the unknown vector case could be defined as in Section 3.1.1.

3.2.2. The elimination approach

We will generalize the approach proposed in [15] and follow the lines of the second elimination approach for the unknown vector case.

For $v_n \in \mathbb{R}^m$, $[v_n]_s$ will denote the $m \times s$ matrix with columns $v_n, \ldots, v_{n+s-1}$. From (34), we have

$$ \Delta x_n = A_1 \Delta z_n^1 + A_2 \Delta z_n^2 + \cdots $$

and

$$ [\Delta x_n]_{q_1} = A_1 [\Delta z_n^1]_{q_1} + A_2 [\Delta z_n^2]_{q_1} + \cdots $$

Assuming that $[\Delta z_n^1]_{q_1}$ (which is a square $q_1 \times q_1$ matrix) is nonsingular, it holds

$$ A_1 = [\Delta x_n]_{q_1} ([\Delta z_n^1]_{q_1})^{-1} - A_2 [\Delta z_n^2]_{q_1} ([\Delta z_n^1]_{q_1})^{-1} - \cdots $$

Thus, A_1 can be replaced in (34) by this expression and we get

$$ x_n - [\Delta x_n]_{q_1} ([\Delta z_n^1]_{q_1})^{-1} z_n^1 = x + A_2 \left\{ z_n^2 - [\Delta z_n^2]_{q_1} ([\Delta z_n^1]_{q_1})^{-1} z_n^1 \right\} + \cdots $$

Setting $E_0(v_n) = v_n$ and defining the vectors $E_k(v_n)$ by

$$E_k(v_n) = E_{k-1}(v_n) - [\Delta E_{k-1}(v_n)]_{q_k} \left([\Delta E_{k-1}(z_n^k)]_{q_k}\right)^{-1} E_{k-1}(z_n^k) \tag{37}$$

it follows, for $k = 1$,

$$E_k(x_n) = x + A_{k+1} E_k(z_n^{k+1}) + A_{k+2} E_k(z_n^{k+2}) + \cdots \tag{38}$$

The transformation is defined by ${}_{11}E_k^{(n)} = E_k(x_n)$. In this transformation, the auxiliary vectors $G_{k,i}^{(n)} = E_k(z_n^i)$ are also needed. They are calculated by replacing v_n by z_n^i in (37). Notice that, for $i \leqslant k$, $G_{k,i}^{(n)} = E_k(z_n^i) = 0$ as in the scalar and unknown vector cases.

Repeating again the same procedure leads to (37) and (38) for any k since $E_k(x) = x$. Thus, the sequence transformation implemented by this recursive algorithm has a kernel given by (34).

This approach could be generalized by considering matrices with $s_k \geqslant q_k$ columns and rank q_k, and by replacing the inverse in (37) by the pseudo-inverse. The other elimination approaches of Section 3.1.2 could also be extended.

3.2.3. The annihilation operator approach

We set $E_0(v_n) = v_n$ and we assume that

$$E_{k-1}(x_n) = E_{k-1}(x) + A_k E_{k-1}(z_n^k) + A_{k+1} E_{k-1}(z_n^{k+1}) + \cdots,$$

which is satisfied for $k = 1$. With the same notation as before, we get

$$[E_{k-1}(x_n)]_{q_k} = [E_{k-1}(x)]_{q_k} + A_k [E_{k-1}(z_n^k)]_{q_k} + A_{k+1} [E_{k-1}(z_n^{k+1})]_{q_k} + \cdots$$

We multiply this relation on the right by $([E_{k-1}(z_n^k)]_{q_k})^{-1}$ and we apply the operator Δ, assuming the nonsingularity of the matrices whose inverses are used. Thus, we obtain

$$\begin{aligned}
&\Delta\left([E_{k-1}(x_n)]_{q_k}\left([E_{k-1}(z_n^k)]_{q_k}\right)^{-1}\right)\\
&\quad= [E_{k-1}(x)]_{q_k}\Delta\left([E_{k-1}(z_n^k)]_{q_k}\right)^{-1}\\
&\qquad+ A_{k+1}\Delta\left([E_{k-1}(z_n^{k+1})]_{q_k}\left([E_{k-1}(z_n^k)]_{q_k}\right)^{-1}\right) + \cdots,
\end{aligned}$$

and, multiplying on the right by the inverse of $(\Delta[E_{k-1}(z_n^k)]_{q_k}^{-1})$, it follows

$$\begin{aligned}
&\Delta\left([E_{k-1}(x_n)]_{q_k}\left([E_{k-1}(z_n^k)]_{q_k}\right)^{-1}\right)\left(\Delta\left([E_{k-1}(z_n^k)]_{q_k}\right)^{-1}\right)^{-1}\\
&\quad= [E_{k-1}(x)]_{q_k} + A_{k+1}\Delta\left([E_{k-1}(z_n^{k+1})]_{q_k}\left([E_{k-1}(z_n^k)]_{q_k}\right)^{-1}\right)\\
&\qquad\qquad\times\left(\Delta\left([E_{k-1}(z_n^k)]_{q_k}\right)^{-1}\right)^{-1} + \cdots
\end{aligned} \tag{39}$$

All the products appearing in this relation have the same general structure, namely

$$Z = (YB^{-1} - XA^{-1})(B^{-1} - A^{-1})^{-1},$$

where A, B, X, Y are matrices. Using the identity

$$(B^{-1} - A^{-1})^{-1} = B - B(B-A)^{-1}B,$$

we obtain

$$\begin{aligned} Z &= Y - Y(B-A)^{-1}B - XA^{-1}B + XA^{-1}B(B-A)^{-1}B \\ &= X + (Y-X) - [Y + XA^{-1}(B-A) - XA^{-1}B](B-A)^{-1}B \\ &= X + (Y-X) - [Y-X](B-A)^{-1}B \\ &= X + (Y-X)[I - (B-A)^{-1}B] \\ &= X + (Y-X)(B-A)^{-1}[(B-A) - B] \\ &= X - (Y-X)(B-A)^{-1}A. \end{aligned}$$

Thus (39) can be written as

$$\begin{aligned} &\left\{[E_{k-1}(x_n)]_{q_k} - \Delta[E_{k-1}(x_n)]_{q_k}\left(\Delta[E_{k-1}(z_n^k)]_{q_k}\right)^{-1}[E_{k-1}(z_n^k)]_{q_k}\right\} \\ &\quad = [E_{k-1}(x)]_{q_k} + A_{k+1}[E_{k-1}(z_n^{k+1})]_{q_k} - \Delta[E_{k-1}(z_n^{k+1})]_{q_k} \\ &\qquad \times \left(\Delta[E_{k-1}(z_n^k)]_{q_k}\right)^{-1}[E_{k-1}(z_n^k)]_{q_k} + \cdots \end{aligned} \tag{40}$$

Let us define $E_k(v_n)$ by

$$E_k(v_n) = E_{k-1}(v_n) - \Delta[E_{k-1}(v_n)]_{q_k}\left(\Delta[E_{k-1}(z_n^k)]_{q_k}\right)^{-1}E_{k-1}(z_n^k). \tag{41}$$

Since $E_{k-1}(v_n)$ is the first column of the matrix $[E_{k-1}(v_n)]_{q_k}$ and since $E_k(x) = E_{k-1}(x)$ and $E_0(x) = x$, then, keeping only the first column of each term in (40), leads to

$$E_k(x_n) = x + A_{k+1}E_k(z_n^{k+1}) + A_{k+2}E_k(z_n^{k+2}) + \cdots$$

Formula (41) is identical to (37). Thus, the annihilation operator approach and the elimination approach lead to the same transformation and the kernel is given by (34).

In the unknown matrix case, the three approaches lead to the same transformation (with $m_k = p_k$ in the direct approach). Moreover, we have, by construction, the following result whose proof is similar to the proofs given in the unknown vector case.

Theorem 4. *If*

$$x_n = x + A_1 z_n^1 + A_2 z_n^2 + \cdots, \quad n = 0, 1, \ldots,$$

with $A_i \in \mathbb{R}^{p\times q_i}$, and if all matrices whose inverse is needed are nonsingular, then, for $k = 0, 1, \ldots,$ it holds, for the transformations $i = 10$ and 11,

$${}_iE_k^{(n)} = x + A_{k+1}G_{k,k+1}^{(n)} + A_{k+2}G_{k,k+2}^{(n)} + \cdots, \quad n = 0, 1, \ldots,$$

where the matrices $G_{k,j}^{(n)}$ are obtained by replacing x_n by z_n^j in (36) for $i = 10$ and v_n by z_n^j in (37) for $i = 11$.

4. Conclusions

In this paper, we showed how to construct new vector sequence transformations. Many issues remain to be studied. First, recursive algorithms for the implementation of the transformations obtained by the direct approaches have to be found. Then, convergence and acceleration results need to be proved. Numerical experiments have to be carried out to see the effectiveness of the transformations. They also have to be compared with the other vector sequence transformations which exist in the literature. Finally, applications have to be studied, in particular to sequences of vectors obtained by iterative methods for the solution of systems of linear and nonlinear equations.

References

[1] T. Ando, Generalized Schur complements, Linear Algebra Appl. 27 (1979) 173–186.
[2] A. Ben-Israel, A note on partitioned matrices and equations, SIAM Rev. 11 (1969) 247–250.
[3] P. Bhimasankaram, On generalized inverses of partitioned matrices, Sankhyā Ser. A 33 (1971) 311–314.
[4] Å. Björck, Numerical Methods for Least Squares Problems, SIAM, Philadelphia, 1996.
[5] C. Brezinski, A general extrapolation algorithm, Numer. Math. 35 (1980) 175–187.
[6] C. Brezinski, Some determinantal identities in a vector space, with applications, in: H. Werner, H.J. Bünger (Eds.), Padé Approximation and its Applications: Bad Honnef 1983, LNM 1071, Springer-Verlag, Berlin, 1984, pp. 1–11.
[7] C. Brezinski, Other manifestations of the Schur complement, Linear Algebra Appl. 111 (1988) 231–247.
[8] C. Brezinski, Vector sequence transformations: methodology and applications to linear systems, J. Comput. Appl. Math. 98 (1998) 149–175.
[9] C. Brezinski, Biorthogonal vector sequence transformations and Padé approximation of vector series, Appl. Numer. Math. 41 (2002) 437–442.
[10] C. Brezinski, M. Redivo Zaglia, Extrapolation Methods: Theory and Practice, North-Holland, Amsterdam, 1991.
[11] C. Brezinski, M. Redivo Zaglia, A general extrapolation algorithm revisited, Adv. Comput. Math. 2 (1994) 461–477.
[12] C. Brezinski, M. Redivo Zaglia, Vector and matrix sequence transformations based on biorthogonality, Appl. Numer. Math. 21 (1996) 353–373.
[13] C. Brezinski, M. Redivo Zaglia, A Schur complement approach to a general extrapolation algorithm, Linear Algebra Appl. 368 (2003) 279–301.
[14] C. Brezinski, H. Sadok, Vector sequence transformations and fixed point methods, in: C. Taylor et al. (Eds.), Numerical Methods in Laminar and Turbulent Flows, Pineridge Press, Swansea, 1987, pp. 3–11.
[15] C. Brezinski, A. Salam, Matrix and vector sequence transformations revisited, Proc. Edinburgh Math. Soc. 38 (1995) 495–510.
[16] C. Brezinski, J. Van Iseghem, Padé approximations, in: P.G. Ciarlet, J.L. Lions (Eds.), Handbook of Numerical Analysis, vol. III, North-Holland, Amsterdam, 1994, pp. 47–222.
[17] C. Brezinski, J. Van Iseghem, A taste of Padé approximation, in: A. Iserles (Ed.), Acta Numerica 1995, Cambridge University Press, Cambridge, 1995, pp. 53–103.
[18] F. Burns, D. Carlson, E. Haynsworth, T. Markham, Generalized inverse formulas using the Schur complement, SIAM J. Appl. Math. 26 (1974) 254–259.

[19] C.A. Butler, T.D. Morley, Six generalized Schur complements, Linear Algebra Appl. 106 (1988) 259–269.

[20] S.L. Campbell, C.D. Meyer Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

[21] D. Carlson, What are Schur complements, anyway?, Linear Algebra Appl. 74 (1986) 257–275.

[22] D. Carlson, E. Haynsworth, T. Markham, A generalization of the Schur complement by means of the Moore–Penrose inverse, SIAM J. Appl. Math. 26 (1974) 169–175.

[23] R.W. Cottle, Manifestations of the Schur complement, Linear Algebra Appl. 8 (1974) 189–211.

[24] R.W. Cottle, On manifestations of the Schur complement, Rend. Sem. Mat. Fis. Milano XLV (1975) 33–40.

[25] M. Fiedler, Special Matrices and their Applications in Numerical Mathematics, Martinus Nijhoff Publishers, Dordrecht, 1986.

[26] M. Gasca, G. Mühlbach, Generalized Schur complements and a test for strict total positivity, Appl. Numer. Math. 3 (1987) 215–232.

[27] T.N.E. Greville, Note on the generalized inverse of a matrix product, SIAM Rev. 8 (1966) 518–521.

[28] E.V. Haynsworth, Reduction of a matrix using properties of the Schur complement, Linear Algebra Appl. 8 (1970) 23–29.

[29] A. Heyting, Die theorie der linearen Gleichungen in einer Zahlenspezies mit nichtkommutativer Multiplikation, Math. Ann. 98 (1927) 465–490.

[30] H.B. Keller, The bordering algorithm and path following near singular points of higher nullity, SIAM J. Sci. Stat. Comput. 4 (1983) 573–582.

[31] K. Jbilou, Méthodes d'Extrapolation et de Projection. Applications aux Suites de Vecteurs, Thèse de 3ième cycle, Université des Sciences et Techniques de Lille, 1988.

[32] G. Marsaglia, G.P.H. Styan, Equalities and inequalities for ranks of matrices, Linear and Multilinear Algebra 2 (1974) 269–292.

[33] G. Marsaglia, G.P.H. Styan, Rank conditions for generalized inverses of partitioned matrices, Sankhyā Ser. A 36 (1974) 437–442.

[34] D.V. Ouellette, Schur complements and statistics, Linear Algebra Appl. 36 (1981) 187–295.

[35] M. Redivo-Zaglia, Pseudo-Schur complements and their properties, Appl. Numer. Math., to appear.

[36] C.A. Rohde, Generalized inverses of partitioned matrices, J. Soc. Indust. Appl. Math. 13 (1965) 1033–1035.

[37] A. Salam, Extrapolation: Extension et Nouveaux Résultats, Thèse, Université des Sciences et Technologies de Lille, 1993.

[38] A. Salam, Non-commutative extrapolation algorithms, Numer. Algorithms 7 (1994) 225–251.

[39] A. Salam, On the vector-valued Padé approximants and the vector ε-algorithm, in: A. Cuyt (Ed.), Nonlinear Numerical Methods and Rational Approximation, II, Kluwer, Dordrecht, 1994, pp. 291–301.

[40] I. Schur, Potenzreihen im Innern des Einheitskreises, J. Reine Angew. Math. 147 (1917) 205–232.

[41] J. Wimp, Sequence Transformations and Their Applications, Academic Press, New York, 1981.

[42] F.-Z. Zhang, Matrix Theory: Basic Results and Techniques, Springer-Verlag, New York, 1999.